Know the Earth...Show the Way

Metadata Monthly

NGA-DCGS Metadata Harmonization

Issue 40
January 2011

TCPED Domain Ontology

Ontology is a model of an information environment consisting of an agreed upon vocabulary, taxonomy, and thesaurus. Domain ontology models a specific area of interest in an information environment. This article discusses the development of domain ontology for a specific and common area of interest in the NSG information environment: the Tasking-Collection-Processing-Exploitation-Dissemination (TCPED) business model.

Why Build the TCPED Domain Ontology?

The driving reason for developing this TCPED domain ontology is to better define and manage data architectures within the National System for Geospatial Intelligence (NSG). This TCPED domain ontology would become one of a collection of semantic structures created to manage the NSG data architecture at the enterprise level. Thus, this TCPED ontology becomes one of many components describing the NSG data architecture.

There are numerous benefits and opportunities to be gained by creating the TCPED domain ontology, including:

- Lifting the important aspects and characteristics of the NSG data environment from many disparate information systems and formally registering and maintaining them to promote a common understanding of the data, facilitating data interoperability and reuse.
- Reusing data and data structures is encouraged by providing the ability to instantiate data and structure from a registry as opposed to hard-coding individual system software components.
- Extending software modules dynamically by not hard-coding data and structure as collections of data and structure mature over time reducing development costs associated with revisiting completed software modules.
- Assembling and maintaining TCPED ontology provides important guidance for those new to the TCPED domain, supporting systems, and knowledge base by communicating the various entities that are used to affect each component of the TCPED business model.
- Promoting data interoperability through a somewhat abstract ontology enables the ontology to be used across data environments and intelligence disciplines.
- Referencing taxonomies and controlled vocabularies associated with this ontology rather than incorporating components into the ontology allows flexibility in definition and control to remain with the originator.

TCPED Domain Classes & Subclasses

Ontology development is a continuous effort, extending as necessary to address fact-of-life changes in an information environment. The initial activity defines the TCPED domain classes and the relationships between those classes. The TCPED domain classes are depicted in Figure 1. Definitions for the TCPED classes were developed in consultation with Chairman Joint Chiefs of Staff (CJCS) Joint Publication 2-01, *Joint and National Intelligence Support to Military Operations.*[1] The TCPED classes identified are:

Tasking: The information analysis to determine the appropriate approach for addressing a specified intelligence problem.

Collection: The collaboration of one or more requirements, systems, and information assets to assemble the necessary data to address a specified intelligence problem.

Processing: The application of one or more data manipulation routines or algorithms to reveal aspects of collected data to assist in drawing conclusions in support of addressing a specified intelligence problem.

Exploitation: The collaboration of intelligence products, analysis capabilities and subject matter expertise to address a specified intelligence problem.

Dissemination: The collaboration of one or more systems, requirements, and mechanisms to provide intelligence information, guidance, and decision support to those tasked with addressing a specified intelligence problem.

[1] CJCS Joint Publication 2-01, *Joint and National Intelligence Support to Military Operations*, 07 October 2004.

Know the Earth...Show the Way
Metadata Monthly
NGA-DCGS Metadata Harmonization
Issue 40
January 2011

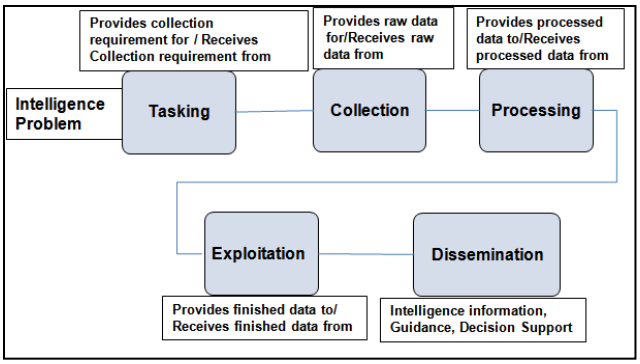

Figure 1. TCPED Domain Classes

Subsequent analysis will define detailed subclasses that semantically describe the domain classes. Each domain class might have one or more subclasses related to it. The thought here is to describe the who, what, why, when, and where aspects of the classes as applicable. Then, for those subclasses, controlled vocabularies or taxonomies will be defined referencing data currently residing in the NSG data environment to populate these structures. As these various semantic structures are defined and reviewed for completeness and accuracy, they will be registered in the DoD Metadata Registry (MDR) in formats intended for both human and machine consumption.

The initial list of TCPED domain ontology subclasses include *Resource*, *Controlled Vocabulary*, *Actor*, and *Topic*. The *Resource* subclass represents the various information sources collaborating to support the instantiation of a TCPED class. As an example, a Tasking requirement can materialize from the collaboration of existing intelligence documents, imagery libraries, and intelligence products, potentially spanning intelligence disciplines. In the NGA Geospatial-Intelligence Information Management System (GIMS), a Tasking requirement takes the form of a Geospatial-Intelligence Information Need (GIN). Other examples include the following requirement types also in the GIMS environment: Processing Requirement (PR), Collection Requirement (CR), Exploitation Requirement (ER), and Dissemination Requirement (DR). Other environments, such as commercial imagery providers, have their own concepts that support the TCPED concept and should be harmonized with this ontology. Other examples of Resources include imagery platforms, sensors, systems, and automated tools that support collection, exploitation and dissemination in the NSG.

The *Controlled Vocabulary* is a list of terms that has been enumerated explicitly. This list is controlled by and is available from a controlled vocabulary registration authority.[2] This is the basis for including controlled vocabularies in the TCPED ontology. Controlled vocabularies are used throughout the NSG data environment. Unfortunately, the controlled vocabularies are typically buried in a database or software module. Identifying, registering, and managing these vocabularies at the NSG enterprise level provides an important component in information sharing and data interoperability.

An *Actor* is the initiator of a TCPED class. This could involve both human and machine-based entities. For example, one human-based entity is an intelligence analyst utilizing a computer-based capability to submit an intelligence requirement. Another example is an automated process that instantiates a CR or ER based on some collection of rule-based criteria.

Finally, the *Topic* subclass represents an association of a TCPED class with one or more semantic information structures or entities for the purpose of associating one or more contexts related to an instantiated TCPED class. These semantic structures can be developed to support any number of contexts or subject areas.

Summary

Unnecessary, redundant duplication and obfuscation can be avoided when a community agrees to document its ontology. Considerable benefits can be enjoyed by all actors in a community when language is plain. However, the most obvious course is not always the easiest course to pursue.

For comments or questions contact:

Pat Gelmetti, *NGA/ASXT (*703-755-5648*)*
Patrick.J.Gelmetti@nga.mil

Dr. Lou York, MITRE (781-271-3880)
lyork@mitre.org

[2] ANSI/NISO Z39.19-2005 ISBN: 1-880124-65-3